The Demand for Food Quality in Rural China

Xiaohua Yu
Ph.D. Candidate
Agricultural, Environmental & Regional Economics and Demography
Penn State University
308 Armsby Building
University Park, PA 16802
Email: xuy103@psu.edu

David Abler
Professor
Agricultural, Environmental & Regional Economics and Demography
Penn State University
207 Armsby Building
University Park, PA 16802
Phone: +1.814.863.8630
Fax: +1.814.865.3746
Email: d-abler@psu.edu

The Demand for Food Quality in Rural China

Rapid economic growth usually leads to significant structural changes in food demand. China is a good case in point.   Since China launched market-oriented economic reforms in 1978, its economic growth rate has averaged 7-8% per year.   Engel's Law predicts that the share of food in total expenditure should decrease as income increases, and China is no exception.   In rural China, the share of food in total expenditure has fallen significantly in recent years, from about 59% in 1994 to about 46% in 2003 (Table 1).   The decline in urban areas in China was also significant during this period, from about 50% to about 37%. Furthermore, as incomes increase, consumers in developing countries tend to shift from less expensive foods such as grains to more expensive foods such as meat and dairy products.   Sahn (1988) found this type of shift in Sri Lanka, as did Ye and Taylor (1995) in rural areas of northern China.   Statistics for rural China as a whole indicate that the share of grain in total food expenditures fell from about 36% in 1994 to about 23% in 2003, while the share of meat in total food expenditures increased from about 17% to about 21% during this period (Table 1).

A number of empirical studies for rural and urban China have analyzed changes in recent years in food demand (e.g., Fan, Wailes and Cramer 1995; Gao, Wailes and Cramer 1996; Huang and Rozelle 1998; Shen 2001; Ma, Rae, Huang and Rozelle 2004; Gould and Dong 2004; Yen, Fang and Su 2004; Wan 2005).   However, a common point in most of the studies for rural China is that the price data are not actual prices but unit values, obtained by dividing expenditures by the quantity consumed.   Relying on unit values can bias empirical analyses because they are not exogenous market prices; they instead reflect household food quality choices within each food product category (Deaton 1988; Nelson 1991).   For example, within the category of "meat" there is considerable scope for household choice with respect to the type

of meat, cut, appearance, texture, tenderness, flavor, nutrient content, freshness, and ease of preparation.

Regressions of quantities demanded on unit values and income may produce biased estimates of income and price elasticities of demand. As we show in this article, the income elasticity is likely to be biased upward, while the absolute value of own-price elasticity is likely to be biased upward for a normal good and downward for an inferior good. The magnitudes of biases can vary but projections of future food and agricultural consumption based on elasticities that do not account for quality could be subject to significant error. In the case of India, Subramanian and Deaton (1996) found that the elasticity of caloric intake with respect to income is about half the income elasticity of total food consumption, due largely to shifts by consumers toward more expensive food groups (such as meat) as income increases but also to shifts toward more expensive foods within each group. The potential for bias has been known for several decades, going back to seminal work by Houthakker (1952), Theil (1952), Prais and Houthakker (1971), and Cramer (1973).

Gale and Huang (2007) analyzed the impact of changes in income on the demand for food quality for several food groups in China using the Prais and Houthakker (1971) methodology and found significant impacts in several cases, including seafood, fruits and vegetables. This methodology starts with the identity $e \equiv pq$, where e represents expenditures on some food group, q is the physical quantity demanded of that food group, and p is the unit price, which is an indicator of quality. Since $p = e/q$, the effect of income on the demand for food quality can be measured by the difference between the elasticity of e respect to income and the elasticity of q with respect to income. Gale and Huang (2007) did not examine the effects of variables other than income on the demand for food quality.

This article has two objectives. The first is to lay out a theoretical framework for assessing the magnitude of bias in estimates of income and price elasticities of demand from studies using unit values that do not account for household food quality choices, and for correcting these biases. The second objective is to estimate the determinants of changes in the quality of food demanded in rural China using panel data for 10 years (1994 through 2003) for rural areas of 26 Chinese provinces.[1] This is a particularly important issue for China given its large and growing role in global food and agricultural markets, and rural China accounts for nearly 60% of China's total population. We analyze nine food products (grains, fats & edible oils, meat, seafood, fresh vegetables, sugar, alcohol, fruits, and dairy products) that account for more than two-thirds of total food expenditures in rural China.

Income and the Demand for Quantity and Quality

Household surveys of consumption typically ask respondents to report their consumption quantities and expenditures for groups of items rather than for specific individual items. In some surveys there are a small number of broadly defined groups (e.g. "meat"); in other surveys there are a larger number of more narrowly defined groups (e.g. "beef," "pork," "poultry," etc.). Even when groups are narrowly defined in the survey, as is the case for rural China, statistical authorities may limit the public release of data to more broadly defined groups. Whether broad or narrow, though, the survey data present us with goods assigned to pre-defined groups and typically without data on prices of specific items within each group. Instead we often have only unit values for each group. The challenge is to analyze the data in a way that makes sense given the economics of the underlying consumer choices with respect to individual items.

Following Deaton (1988), consider a food group i composed of a number of items, such as different types of meat or different varieties of grain. Let $\underset{\sim}{p}_{ij}$ denote the vector of prices for the items in this food group in some region j. Assume that these prices can be written as the product of a common (or average) vector of prices for all regions ($\underset{\sim}{p}_i^*$) and a term reflecting differences between regions in prices (λ_{ij}):

$$\underset{\sim}{p}_{ij} = \lambda_{ij}\underset{\sim}{p}_i^* \tag{1}$$

Inter-regional differences in prices could arise due to transportation costs or, perhaps in some countries, government price, tax, or distribution policies. λ_{ij} is typically assumed in models of this type to be exogenous to household consumption decisions and we generally maintain that assumption here.

The aggregate quantity of food group i consumed in region j (Q_{ij}) as typically measured in household surveys is simply the total number of kilograms or pounds of all the items within that food group:

$$Q_{ij} = \underset{\sim}{\theta}_i'\underset{\sim}{q}_{ij}, \tag{2}$$

where $\underset{\sim}{q}_{ij}$ is a vector of consumption quantities (measured by weight) for the items within this food group and $\underset{\sim}{\theta}_i$ is a vector of ones.[2] For other purposes, such as nutritional studies, $\underset{\sim}{\theta}_i$ might contain information on the caloric, protein, or other nutrient content of food items. Expenditures on food group i in region j (E_{ij}) are:

$$E_{ij} = \underset{\sim}{p}_{ij}'\underset{\sim}{q}_{ij}. \tag{3}$$

Given this, the unit value of food group i in region j (V_{ij}) is:

$$V_{ij} = \frac{E_{ij}}{Q_{ij}} = \frac{\underset{\sim}{p}_{ij}' \underset{\sim}{q}_{ij}}{\underset{\sim}{\theta}_i' \underset{\sim}{q}_{ij}} = \lambda_{ij} \left(\frac{\underset{\sim}{p}_i^{*'} \underset{\sim}{q}_{ij}}{\underset{\sim}{\theta}_i' \underset{\sim}{q}_{ij}} \right) = \lambda_{ij} \upsilon_{ij} ,$$ (4)

where $\upsilon_{ij} = \underset{\sim}{p}_i^{*'} \underset{\sim}{q}_{ij} / \underset{\sim}{\theta}_i' \underset{\sim}{q}_{ij}$ is a measure of quality. υ_{ij} is the average cost of food items within

group i consumed in region j, controlling for inter-regional differences in prices. It is

endogenous because it depends on household food consumption choices, which in turn depend

on income, prices, and household characteristics. Equation (4) implies that

$$\ln V_{ij} = \ln \lambda_{ij} + \ln \upsilon_{ij} ,$$ (5)

which can be viewed as a hedonic model of unit values for food groups.

Following Deaton (1988) and Deaton and Muellbauer (1980), assume that the demand for

food group i is weakly separable from all other food and non-food groups, so that we have a

two-stage budgeting problem where consumers in the first stage choose how much to spend on

each group and in the second stage decide how to allocate expenditures for each group among

the goods in that group. The utility function in a two-stage budgeting problem can be written as

$u_j = u_j \left(a_{1j}, a_{2j}, \ldots, a_{nj} \right)$, where $a_{ij} = a_{ij} \left(\underset{\sim}{q}_{ij} \right)$ is an aggregate of goods within the ith group in

region j and n is the total number of groups. The utility maximization process in a

two-stage budgeting problem yields a vector of group price indices ($\underset{\sim}{\pi}_j$), with π_{ij} equal to the

marginal cost of a_{ij}. The group price indices are endogenous shadow prices.[3]

Optimal expenditures on group i at the first stage of the two-stage budgeting problem

are in general a function of prices of all goods in all groups (denoted by the vector $\underset{\sim}{P}_j$), total

income (Y_j), and a vector of other household characteristics affecting consumption ($\underset{\sim}{Z}_j$):

$$E_{ij} = g_{ij} \left(\underset{\sim}{P}_j, Y_j, \underset{\sim}{Z}_j \right) .$$ (6)

At the second stage, optimal demands within food group i are a function of prices of goods within that group, group expenditures, and household characteristics:

$$q_{ij} = f_{ij}\left(p_{ij}, E_{ij}, Z_j\right) = f_{ij}\left(p_i^*, E_{ij}/\lambda_{ij}, Z_j\right). \tag{7}$$

The second equality in (7) follows from the fact that the demand functions are homogenous of degree zero in group expenditures and prices. Equations (2), (6), and (7) imply that the demand for the aggregate quantity of food group i is

$$Q_{ij} = \theta_i' f_{ij}\left(p_{ij}, E_{ij}, Z_j\right) = \theta_i' f_{ij}\left(p_{ij}, g_{ij}\left(P_j, Y_j, Z_j\right), Z_j\right) = h_{ij}\left(p_{ij}, g_{ij}\left(P_j, Y_j, Z_j\right), Z_j\right), \tag{8}$$

where the group-level function $h_{ij}(\cdot)$ aggregates the information from the vector of product-specific functions $f_{ij}(\cdot)$ within that group.

However, equation (8) often cannot be empirically estimated because data on prices of individual items are unavailable, as is the case for rural China. Instead researchers typically replace the vector of prices p_{ij} by the unit value V_{ij} and the vector of all prices P_j by the corresponding vector of unit values V_j to obtain an equation that can be estimated:

$$Q_{ij} \approx h_{ij}\left(V_{ij}, g_{ij}\left(V_j, Y_j, Z_j\right), Z_j\right), \tag{9}$$

Note that $V_{ij} = \pi_{ij}$ only if the aggregator for group i in the utility function is identical to equation (2): $a_{ij}\left(q_{ij}\right) = Q_{ij} = \theta_i' q_{ij}$. If the vector θ_i consists of ones, so that food items are aggregated by weight, this would imply that consumers have no interest in quality differences within a food group. In this case consumers would purchase only the least expensive item within each group and spend nothing on the other items, leaving V_{ij} equal to the lowest price in the vector of prices p_{ij}. This is the only situation in which the replacement of p_{ij} by V_{ij} and

the replacement of $\underset{\sim}{P_j}$ by $\underset{\sim}{V_j}$ can be justified theoretically, because the only price in each group that affects consumer decision-making is the lowest price.[4]

In econometric work, estimating the parameters of equation (9) to examine the effects of a change in income on expenditures and in turn Q_{ij} implicitly involves holding V_{ij} constant as Y_j changes. However, unit values cannot stay constant when income changes unless income has no impact on the quality of goods purchased within each group (υ_{ij}) or there is an offsetting change in the inter-regional price factors (λ_{ij}) that leaves unit values unchanged.

Let $\gamma_{ij} = \left(\partial \ln Q_{ij} / \partial \ln E_{ij}\right)\left(d \ln g_{ij} / d \ln Y_j\right)$ denote the income elasticity of demand for food group i as obtained from equation (9), and let $\eta_{ij} = d \ln \upsilon_{ij} / d \ln Y_j$ denote the elasticity of demand for quality within group i with respect to income, which we would typically presume is positive. Consider what happens if there is an offsetting change in λ_{ij} that leaves V_{ij} unchanged, so that $d \ln \lambda_{ij} / d \ln Y_j = -\eta_{ij}$. In an econometric analysis this would be tantamount to relying on inter-regional price variability to reduce collinearity between Y_j and V_{ij} to the point where equation (9) could be reliably estimated. Let $\varepsilon_{ij} = -d \ln Q_{ij} / d \ln \lambda_{ij}$ be minus one times the elasticity of food consumption with respect to the inter-regional price factor. We would typically presume a downward-sloping demand curve for food ($\varepsilon_{ij} > 0$). Utilizing equation (9),

$$\tilde{\gamma}_{ij} = \left.\frac{d \ln Q_{ij}}{d \ln Y_j}\right|_{\lambda_{ij}\text{ offsetting}} = \left.\frac{d \ln Q_{ij}}{d \ln Y_j}\right|_{\lambda_{ij}\text{ constant}} + \frac{d \ln Q_{ij}}{d \ln \lambda_{ij}}\frac{d \ln \lambda_{ij}}{d \ln Y_j} = \gamma_{ij} + \varepsilon_{ij}\eta_{ij}. \tag{10}$$

If $\eta_{ij} > 0$ and $\varepsilon_{ij} > 0$, equation (10) implies the elasticity of food consumption with respect to income is greater with an offsetting change in λ_{ij} than when λ_{ij} is constant ($\tilde{\gamma}_{ij} > \gamma_{ij}$).

Estimation of equation (9) assuming that unit values are exogenous, as is typically the case, will overstate the responsiveness of consumption to income by the amount $\varepsilon_{ij}\eta_{ij}$. If food group i is normal ($\gamma_{ij} > 0$), γ_{ij} will be closer to 0 than $\tilde{\gamma}_{ij}$. If it is inferior ($\gamma_{ij} < 0$), γ_{ij} will be further away from 0 (more negative) than $\tilde{\gamma}_{ij}$. This is a type of simultaneous equation bias that occurs in estimation because unit values are in fact endogenous. The larger the income elasticity of demand for food quality, the greater the magnitude of bias.

Deaton (1988) indicates that estimation of equation (9) will also tend to overstate the responsiveness of consumption to changes in price, assuming that the product in question is a normal good. Consider a change in prices due to a change in the inter-regional price factor λ_{ij}. The correct value of the own-price elasticity of demand is $-\varepsilon_{ij}$. As Deaton (1988) shows, mistakenly measuring the price elasticity by the relationship between quantity and unit value yields a different elasticity ($\tilde{\varepsilon}_{ij}$):

$$\tilde{\varepsilon}_{ij} = -\frac{d\ln Q_{ij}}{d\ln V_{ij}} = \frac{\varepsilon_{ij}}{1 - \varepsilon_{ij}\eta_{ij}/\gamma_{ij}}. \tag{11}$$

If food group i is normal ($\gamma_{ij} > 0$) and if $0 < 1 - \varepsilon_{ij}\eta_{ij}/\gamma_{ij} < 1$, the absolute value of the own-price elasticity of demand will be overstated ($\tilde{\varepsilon}_{ij} > \varepsilon_{ij}$).[5] The larger the income elasticity of demand for food quality (η_{ij}), the smaller the denominator in equation (11) and the greater the degree of overestimation. The overestimation occurs because an increase in the inter-regional price factor λ_{ij} has a negative income effect in this situation on the demand for food quality, causing V_{ij} to rise by less in percentage terms than λ_{ij}. This makes it appear as if

consumption is more responsive to price when looking at the ratio $d\ln Q_{ij}/d\ln V_{ij}$ than is actually the case.

On the other hand, if food group i is inferior ($\gamma_{ij} < 0$), then $1 - \varepsilon_{ij}\eta_{ij}/\gamma_{ij} > 1$ and the absolute value of the own-price elasticity of demand will be understated ($\tilde{\varepsilon}_{ij} < \varepsilon_{ij}$). The larger the income elasticity of demand for food quality (η_{ij}), the larger the denominator in equation (11) and the greater the degree of underestimation of the absolute value of the own-price elasticity. This occurs because an increase the inter-regional price factor λ_{ij} has a positive income effect in this situation on the demand for food quality, causing V_{ij} to rise by more in percentage terms than λ_{ij} and making it appear as if consumption is less responsive to price when looking at the ratio $d\ln Q_{ij}/d\ln V_{ij}$ than is actually the case.

Equations (10) and (11) predict that the distortions in estimates of income and own-price elasticities of demand from estimation of equation (9) will depend on the correct value of the own-price elasticity. The larger the value of ε_{ij}, the greater the degree of overestimation of the income elasticity and the greater the degree of either overestimation or underestimation of the absolute value of the own-price elasticity, depending on whether product is normal or inferior.

Given values for η_{ij}, $\tilde{\varepsilon}_{ij}$ and $\tilde{\gamma}_{ij}$, equations (10)-(11) can be viewed as a system of two equations in two unknowns, the correct values for the price and income elasticities of demand for quantity (ε_{ij} and γ_{ij}). Viewing the equations in this way is useful because estimates of $\tilde{\varepsilon}_{ij}$ and $\tilde{\gamma}_{ij}$ are available from existing studies of food demand that do not correct for quality, and our study (as well as other studies) provide estimates of η_{ij}. Letting $b_{ij} = \tilde{\gamma}_{ij} - 2\tilde{\varepsilon}_{ij}\eta_{ij}$ and $c_{ij} = \tilde{\varepsilon}_{ij}\eta_{ij}\tilde{\gamma}_{ij}$, the solution for γ_{ij} is:

$$\gamma_{ij} = \frac{b_{ij} + \sqrt{b_{ij}^2 + 4c_{ij}}}{2} \quad \text{if} \quad \tilde{\gamma}_{ij} > 0, \tag{12a}$$

$$\gamma_{ij} = \frac{b_{ij} - \sqrt{b_{ij}^2 + 4c_{ij}}}{2} \quad \text{if} \quad \tilde{\gamma}_{ij} < 0. \tag{12b}$$

With this solution in hand, ε_{ij} can be obtained using equation (10):

$$\varepsilon_{ij} = \frac{\tilde{\gamma}_{ij} - \gamma_{ij}}{\eta_{ij}}. \tag{13}$$

If $\tilde{\gamma}_{ij} = 0$, it can be shown that equations (10)-(11) degenerate to the solution $\varepsilon_{ij} = \gamma_{ij} = 0$.

The solutions for ε_{ij} and γ_{ij} when $\tilde{\gamma}_{ij} > 0$ are illustrated in Figure 1, while the solutions when $\tilde{\gamma}_{ij} < 0$ are shown in Figure 2. Equation (10) in Figure 1 is a downward-sloping straight line between $\tilde{\gamma}_{ij}$ on the *y*-axis and $\tilde{\gamma}_{ij}/\eta_{ij}$ on the *x*-axis, while equation (11) is an increasing function that starts at the origin and has an asymptote at $\tilde{\varepsilon}_{ij}$. In Figure 2 equation (10) is a downward-sloping straight line beginning at $\tilde{\gamma}_{ij}$ on the *y*-axis while equation (11) is an increasing function that has an asymptote at $\tilde{\varepsilon}_{ij}$ from below and 0 from above. Estimates of ε_{ij} and γ_{ij} from equations (12a) or (12b) and (13) can be compared to values of $\tilde{\varepsilon}_{ij}$ and $\tilde{\gamma}_{ij}$ to gauge the degree to which regressions of quantities demanded on unit values and income bias price and income elasticities of demand, and to correct for these biases.

Econometric Model and Data

The econometric model specified here follows in the footsteps of Cox and Wohlgenant (1986) and Deaton (1988), who used cross-sectional data to estimate the determinants of food quality choices in the U.S. and Côte d'Ivoire, respectively. However, in contrast to their

cross-sectional analyses, we use panel data at the provincial level for rural China. Panel data analysis can overcome the difficulty of unobservable variables affecting consumer choices, such as spatial factors, and can improve the efficiency of a regression (Hsiao 2003).

Adding a subscript t to denote time, and assuming that the inter-regional price factors (λ_{ij}) are time-invariant, the empirical counterpart to the hedonic model in equation (5) is

$$\ln V_{ijt} = \ln \lambda_{ij} + \ln \upsilon_{ijt} . \tag{14}$$

Deaton and Muellbauer (1980) developed a panel data model to estimate quality choice along the following lines:

$$\ln \upsilon_{ijt} = \ln p_{it}^* + h\left(Y_{jt}, \underset{\sim}{Z}_{jt}\right), \tag{15}$$

where p_{it}^* is a common reference price for food group i for all regions, similar to the theoretical model above, Y_{jt} is income, and $\underset{\sim}{Z}_{jt}$ is a vector of household characteristics.

In the empirical literature on food demand, household characteristics found to be important include household size, place of residence, and the age, gender, education, race, ethnicity, and employment of household members (Cox and Wohlgenant 1986; Behrman and Deolalikar 1987; Deaton 1988; Dong, Shonkwiler and Capps 1998; Gould and Dong 2004; Ye and Taylor 1995). Household size squared is sometimes also included to test for nonlinearities with respect to this variable (e.g., scale economies at small household sizes and scale diseconomies at large sizes).

Combine equations (14)-(15) and assume that the function $h(\cdot)$ is log-linear. The log-linear form can be viewed as a first-order Taylor series approximation to the true but unknown function and is frequently adopted in empirical hedonic model studies (Deaton and

Muellbauer 1980).　Then, assuming for simplicity that $\ln p_{it}^*$ can be proxied by a time trend

(t), the hedonic model that we estimate is:

$$
\begin{aligned}
\ln V_{ijt} = \beta_{i0} &+ \beta_{i1} \ln PCI_{it} + \beta_{i2} \ln HHSIZE_{it} + \beta_{i3} (\ln HHSIZE_{it})^2 + \beta_{i4} \ln LABOR_{it} \\
&+ \beta_{i5} \ln HOUSE_{it} + \beta_{i6} \ln LAND_{it} + \beta_{i7} \ln EDEXP_{it} + \beta_{i8} \ln EDLEVEL_{it} + \gamma_i t + \mu_{ij} + e_{ijt}
\end{aligned}
\tag{16}
$$

where PCI_{it} is per capita income, $HHSIZE_{it}$ is average household size, $LABOR_{it}$ is the

average number of members of each household who participate in the labor force, $HOUSE_{it}$ is

the average house area (in square meters) per capita, $LAND_{it}$ is average cropland area (in mu)

per capita, $EDEXP_{it}$ is expenditures on education per capita, $EDLEVEL_{it}$ is the fraction of the

adult population with more than a primary school education, μ_{ij} ($= \ln \lambda_{ij}$) is a term reflecting

regional differences, and e_{ijt} is an independently and identically distributed error term.[6]

Equation (16) is a typical panel data model.　The model can be estimated by either fixed

effects or random effects.　If μ_{ij} is a random variable, so that there are no systematic

differences between regions, a random effects model is preferred because it is more efficient than

a fixed effects model.　Otherwise a fixed effects model is superior.　Hausman's (1978)

specification test between fixed and random effects models can be used to analyze whether there

are systematic differences among regions.

The panel dataset consists of data for 10 years (1994 through 2003) for rural areas of 26

Chinese provinces, with data being at the provincial level.　We analyze nine food products

(grains, fats & edible oils, meat, seafood, fresh vegetables, sugar, alcohol, fruits, and dairy

products) that account for more than two-thirds of total food expenditures in rural China.　Data

are from the China National Statistics Bureau (CNSB).　The dataset begins in 1994 in order to

avoid prior years in which prices were significantly distorted by government regulations.　Even

though China began food policy reforms in the late 1970s, price regulations were not abandoned until 1993 (Ma et al. 2004).

Unit values for 1994 are derived from *Rural Household Survey Statistics* (RHSS), a CNSB publication, dividing total expenditure in each food group by the total quantity consumed. Starting with the 1994 unit values, we use the provincial-level consumer price index (CPI) for each food group for 1995 through 2003 to compute unit values for each food group for those years. Provincial-level CPIs are obtained from the *China Statistical Yearbook of Prices and Urban Household Survey* (various editions), published by CNSB. Data for the right-hand side variables in equation (16) are from *Rural Household Survey Statistics* (various editions). *Rural Household Survey Statistics* covers 27 provinces, of which Tibet is excluded from our analyses because of missing data, leaving 26 provinces. Nominal values are converted to real terms using the overall rural China CPI, with all prices expressed in 1994 Yuan.

Hausman Test Results for Spatial Differences

The null hypotheses of the Hausman tests are that there are no systematic differences in unit values and quantity demand among the cross-sectional cohorts—the 26 provinces in our study. Hausman test results are reported in Table 2. The null hypothesis of no systematic differences in demand for quality among regions can be rejected at the 5% significance level for grains, meat, sugar, and alcohol. For these food groups, a fixed effects model is preferred. Grains, meat, and alcohol are products in which until the 1990s there was little inter-provincial or even inter-farm trade in rural China (Huang and Rozelle 1998). Markets for grains have become much more integrated across space since then (Huang and Rozelle 2006). However, rates of commercialization for grains, defined as the ratio of grains purchased to all grains

consumed (purchased plus produced by the household), remain relatively low, which could help explain the Hausman test results for that food group.   Rural China differs from urban China in this regard.   With respect to sugar, there are significant differences in product composition across Chinese provinces that may account for the Hausman test results.

We cannot reject the null hypothesis of no systematic inter-provincial differences in demand for quality for fats & edible oils, vegetables, seafood, and fruits.   These products are highly commercialized and now largely standardized throughout China owing to economic reforms and market development since 1978.   Quality differences among provinces are small, so the random effects model is preferred for these products.

The Hausman test for dairy products is negative, implying that the data fail to satisfy the asymptotic assumptions of the test.   However, the parameter estimates of the random and fixed effects models are very similar to each other.   In the following discussion, the fixed effects results are used for dairy.

Hedonic Model Results and Discussion

The hedonic model results are shown in Table 2.   Overall, the results are good, with reasonable R^2 values for most food groups (vegetables being the exception) and with most explanatory variables statistically significant.

Income

The results indicate that per capita income is statistically significant for five food groups: grains (estimated elasticity of 0.31), fats & oils (0.19), seafood (0.17), vegetables (0.35), and dairy products (0.18).   The estimated income elasticities of demand for quality for the two of these products generally viewed as necessities—grains and vegetables—are larger than those for

fats & oils, seafood, and dairy products, which are generally viewed as luxuries in the case of rural China.   As incomes increase, it appears as if consumers in rural China make greater adjustments to the quality of necessities they consume than the quality of luxuries.

Estimates of own-price and income elasticities of demand from the literature for rural China that do not correct for quality (e.g., Huang and Rozelle 1998; Shen 2001; Ma et al. 2004) can be used in conjunction with the results here and equations (12a), (12b), and (13) to obtain estimates of the quality-corrected price and income elasticities of demand.   The corrected results are shown in Table 3 for grains, vegetables, seafood, and fats & oils.   These are the four food groups for which we have both (1) a statistically significant income elasticity of demand for quality in this study and (2) available estimates from the literature of price and income elasticities of demand for quantity for rural China.   In the case of grains, the results indicate that the income (expenditure) elasticity of demand for grains in the literature is overstated by more than 30% once the demand for quality is taken into account, and the own-price elasticity of demand for grains is overstated in absolute value by more than 40%.[7]   Smaller though still significant biases are found for vegetables and seafood.

Household Size and Labor Force Participation

The estimated coefficients for both the log of household size and the square of this variable are statistically significant for grains, seafood, alcohol and fruits.   The relationship between demand for quality and household size for grains (a necessity) is dome-shaped, implying that as household size increases, the demand for grain quality increases at first and then decreases, with a peak at a household size of about 4.6.   However, the relationship for seafood, alcohol, and fruits (mainly luxury products) are U-shaped, implying that as household size increases the demand for quality decreases at first and then increases; the minima are at

15

household sizes of approximately 5.4, 7.4, and 7.2 respectively.   Considering that the average household size in rural China is 4.45, these results generally imply that there are scale diseconomies with respect to household size in the choice of food quality for luxuries.

Gould and Dong (2004), in a study for urban China using household survey data, found a positive relationship between household size and the demand for quality for pork, seafood and vegetables; a negative relationship for fats & oils and other food products; and no statistically significant relationship for beef, poultry, fruits, rice, other grains, dairy products or eggs.   Their model did not allow for the possibility of a dome-shaped or U-shaped relationship between household size and demand for quality.   Our results differ from hedonic studies of durable goods prices, which generally find scale economies in household size, perhaps in part because food (unlike most durables) is a rival good within the household.

The labor force participation variable is positive and statistically significant for four of the nine products—fats & oils, alcohol, fruits, and dairy products.   It is not statistically significant for the other five products.   A higher rate of labor force participation implies greater current income and also greater permanent income, suggesting that the effects of income on the demand for food quality are not fully captured by the per capita income variable.   Perhaps there is some remaining income effect that is being captured by the labor force participation variable.

House Area and Cropland Area

The estimated coefficients for house area are negative and statistically significant for two products: seafood and dairy products.   Housing can represent a large share of total household expenditures and as such may crowd out food expenditures.

The estimated coefficients for cropland area for grains, sugar, seafood, and dairy products are statistically significant, and among these are positive except for seafood.   As cropland area

increases, households have greater current income and greater permanent income.   As with labor force participation, there may be some remaining effect here of income on the demand for food quality that is not being captured by the per capita income variable.

Education

The estimated coefficients for the educational expenditure variable are negative and statistically significant for five products: grains, meat, vegetables, sugar, and dairy products. The estimated coefficients for the level-of-education variable are negative and statistically significant for grains, positive and statistically significant for sugar, and not statistically significant for the other products.

Households in rural China have made significant investments in education in recent years, and the level of education in rural China has increased rapidly.   The percentage of the adult population with more than a primary school education increased from 28% in 1983 to 63% in 2003.   The share of total household expenditures devoted to education in rural China has risen even more quickly, from 1.8% in 1983 to 9.0% in 2003.   Returns to education are significant but can take many years to materialize.   In such a situation, households may sacrifice short-run interests such as food quality in order to achieve higher incomes in the future through education.   In particular, grains represent a large share of total household expenditures (about 10% in 2003), so it makes sense that education has crowding-out effects on grain quality.

In contrast to our results, Gould and Dong's (2004) study for urban China finds a generally positive relationship between education and the demand for food quality.   Unlike our study, they did not include income as an explanatory variable in their unit value regressions. Education and income are positively correlated, so their results for education may reflect the influence of income.

Households choose educational expenditures contemporaneously with food expenditures, so it is possible that our educational expenditures variable is endogenous. Results of the Hausman (1978) for endogeneity, using educational expenditures lagged one year as an instrumental variable for current educational expenditures, fail to reject the null hypothesis of exogeneity except in one case, the random effects model for fats & oils. And in that case the educational expenditures variable is not statistically significant in either the model reported in Table 2 or in the instrumental variable model.

Changes in Real Food Prices over Time

The time trend variable is negative and statistically significant for eight of the nine products, the only exception being vegetables where it is not statistically significant. Other things held constant, the results suggest that real unit values have been falling over time. China's transition from a planned economy to a market economy has stimulated much productivity-increasing technical change and significantly reduced transaction costs (Huang and Rozelle 1998; Huang and Rozelle 2006), which may help explain these results. Also, some commodities were significantly protected from international competition in the 1990s, particularly wheat and soybeans. Protection for these commodities has been declining, causing domestic prices to move toward world prices (Huang et al. 2007).

Conclusions

The objectives of this article were to develop a theoretical framework for assessing bias in estimates of income and price elasticities of demand in studies using unit values that do not account for household food quality choices; and then to estimate the determinants of changes in the quality of food demanded in rural China, using panel data for 10 years (1994 through 2003)

for rural areas of 26 Chinese provinces.   We analyzed nine food products (grain, fats & edible oils, meat, seafood, fresh vegetables, sugar, alcohol, fruits, and dairy products) that account for more than two-thirds of total food expenditures in rural China.

Our theoretical framework indicates that the income elasticity is likely to be biased upward, while the absolute value of own-price elasticity is likely to be biased upward for a normal good and downward for an inferior good.   The larger the income elasticity of demand for food quality, the greater the degree of bias in both the income and own-price elasticities. Our framework also provides a means for recovering the correct values of the income and price elasticities of demand using estimates of these elasticities from studies of food demand that do not correct for quality and estimates of the income elasticity of demand for food quality.

The theoretical framework is needed because household surveys almost always present us with data in which individual goods have been assigned to pre-defined groups and typically lack data on prices of specific items within each group.   Instead we often have only unit values for each group.   The challenge is to analyze the data in a way that makes sense given the economics of the underlying consumer choices with respect to individual items.   An ideal dataset would not group items at all but would provide the price and quantity for every item consumed.   Some type of grouping would probably still be necessary after the fact to make the econometric analysis tractable, but the group demand functions would be consistent with theory on multistage demand systems, estimable without resort to unit values, and avoid the biases in price and income elasticities of demand found here.

Our econometric results indicate that households in rural China tend to consume higher quality food as income increases, with a greater sensitivity to income for basic foods such as grains than for luxury foods.   These results suggest that existing studies of food demand for

rural China that do not correct for food quality are biased, because as income increases households switch from lower-quality food to higher-quality food.   For grains, our results suggest that the income elasticity of demand for grains in the literature is overstated by more than 30% once the demand for quality is taken into account, and the own-price elasticity of demand for grains is overstated in absolute value by more than 45%.   Smaller but still significant biases are also found for vegetables and seafood.

Our results indicate that there are systematic price differences among provinces mainly for self- and locally-sufficient foods, such as grains and meat, but no systematic price differences for highly commercialized products such as seafood and dairy products. We also find that the pursuit of additional education in rural China has had significant crowding-out effects on the demand for food quality, in particular for grains. Households tend to sacrifice short-run interests by consuming lower-quality grain in order to pursue additional education and achieve higher incomes in the future.   In addition, because of productivity-enhancing technical change and a reduction in transaction costs resulting from economic reforms, real food prices in China have fallen in recent years.

Considering the rapid rate of China's economic growth in recent years and the importance of China to global food and agricultural markets, projections of future food demand for China should take into account the growing demand for food quality.   Failing to do so could lead to overestimates of future growth in the quantity of food consumed in China, missing a shift from simply more food to better quality food.

References

Behrman, J. R., and A. B. Deolalikar. 1987. "Will Developing Country Nutrition Improve with Income? A Case Study for Rural South India." *Journal of Political Economy* 95(3):492-507.

Cox, T. L. and M. K. Wohlgenant. 1986. "Prices and Quality Effects in Cross-Section Demand Analysis." *American Journal of Agricultural Economics* 68(4):908-919.

Cramer, J. S. 1973. "Interaction of Income and Price in Consumer Demand." *International Economic Review* 14(2):351-363.

Deaton, A., and J. Muellbauer. 1980. *Economics and Consumer Behavior*. Cambridge: Cambridge University Press.

Deaton, A. 1988. "Quality, Quantity and Spatial Variation of Price." *American Economic Review* 78(3):418-430.

Dong, D., J. S. Shonkwiler, and O. Capps, Jr. 1998. "Estimation of Demand Functions Using Cross-Sectional Household Data: The Problem Revisited." *American Journal of Agricultural Economics* 80(3):466-473.

Fan, S., E. J. Wailes, and G. L. Cramer. 1995. "Household Demand in Rural China: A Two-Stage LES-AIDS Model." *American Journal of Agricultural Economics* 77(1):54-62.

Gale, F., and K. Huang. 2007. *Demand for Food Quantity and Quality in China*. USDA, Economic Research Service, Economic Research Report No. 32. Washington, DC: Government Printing Office. http://www.ers.usda.gov/publications/err32/err32.pdf.

Gao, X. M., E. J. Wailes, and G. L. Cramer. 1996. "A Two-Stage Rural Household Demand

Analysis: Microdata Evidence from Jiangsu Province, China." *American Journal of*

Agricultural Economics 78(3):604-613.

Gould, B. W., and D. Dong. 2004. "Product Quality and the Demand for Food: The Case of

Urban China." Paper presented at AAEA Annual Meeting, Denver, CO, August 2004.

http://agecon.lib.umn.edu/cgi-bin/pdf_view.pl?paperid=14030.

Hausman, J. A. 1978. "Specification Tests in Econometrics." *Econometrica* 46(6):1251-1271.

Hsiao, C. 2003. *Analysis of Panel Data*, 2nd edition. Cambridge: Cambridge University Press.

Huang, J., Y. Liu, W. Martin, and S. Rozelle. 2007. "Integrating China's Agricultural Economy

into Global Market: Measuring Distortions in China's Agricultural Sector." Paper

presented at International Agricultural Trade Research Consortium Summer Symposium,

Beijing, China, July 2007. http://aede.osu.edu/programs/Anderson/trade/10Huang.pdf.

Huang, J., and S. Rozelle. 2006. "The Emergence of Agricultural Commodity Markets in China."

China Economic Review 17(3):266-280.

Huang, J., and S. Rozelle. 1998. "Market Development and Food Demand in Rural China."

China Economic Review 9(1):25-45.

Houthakker, H. S. 1952. "Compensated Changes in Quantities and Qualities Consumed."

Review of Economic Studies 19(3):55-164.

Ma, H., A. Rae, J. Huang, and S. Rozelle. 2004. "Chinese Animal Product Consumption in

1990s." *Australian Journal of Agricultural and Resources Economics* 48(4):569-590.

Moschini, G. 2001. "A Flexible Multistage Demand System Based on Indirect Separability."

Southern Economic Journal 68(1):22-41.

Nelson, J. A. 1991. "Quality Variation and Quantity Aggregation in Consumer Demand for Food." *American Journal of Agricultural Economics* 73(4):1204-1212.

Prais, S. J., and H. S. Houthakker. 1971. *The Analysis of Family Budgets*, 2nd edition. Cambridge: Cambridge University Press.

Sahn, D. E. 1988. "The Effect of Price and Income Changes on Food-Energy Intake in Sri Lanka." *Economic Development and Cultural Change* 36(2):315-340.

Shen, J. 2001. "Structural Change in Household Demand in Rural China." *Nou Rin Gyou Mon Dai Ken Kyu* (Japanese) 37(2):11-23.

Subramanian, S., and A. Deaton. 1996. "The Demand for Food and Calories." *Journal of Political Economy* 104(1):133-162.

Theil, H. 1952. "Qualities, Prices and Budget Enquiries." *Review of Economic Studies* 19(3):129-147.

Wan, G. H. 2005. "Convergence in Food Consumption in Rural China: Evidence from Household Survey Data." *China Economic Review* 16(1):90-102.

Ye, X., and J. E. Taylor. 1995. "The Impact of Income Growth on Farm Household Nutrient Intake: A Case Study of a Prosperous Rural Area in Northern China." *Economic Development and Cultural Change* 43(4):805-819.

Yen, S. T., C. Fang, and S.-J. Su. 2004. "Household Food Demand in Urban China: A Censored System Approach." *Journal of Comparative Economics* 32(3):564-585.

Figure 1
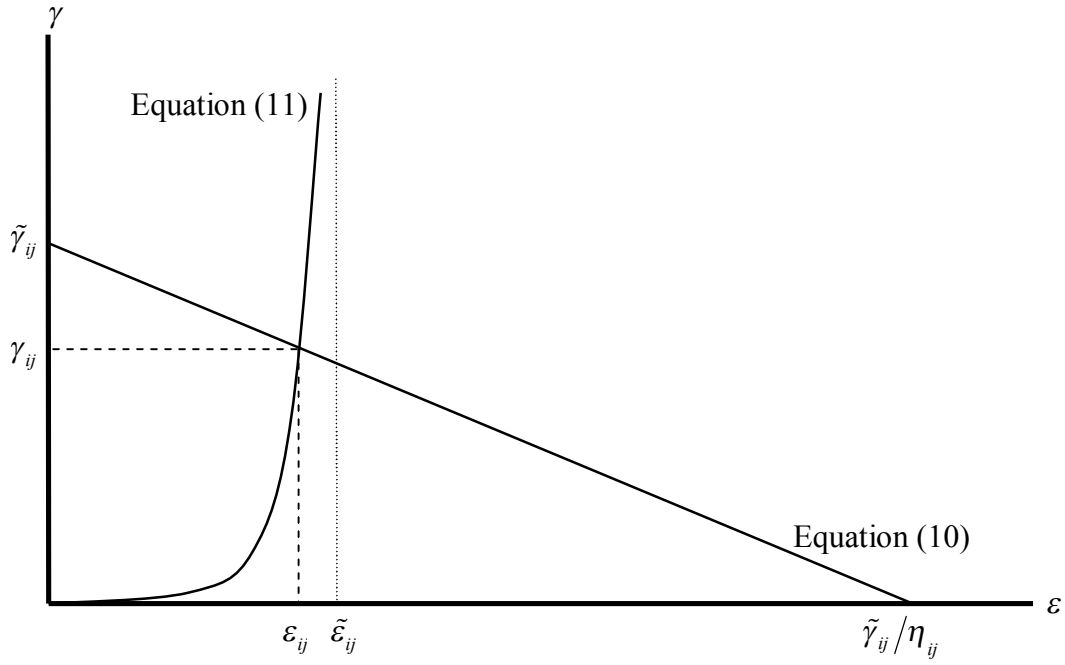**Solutions for Income and Price Elasticities (Normal Good Case)**

Figure 2
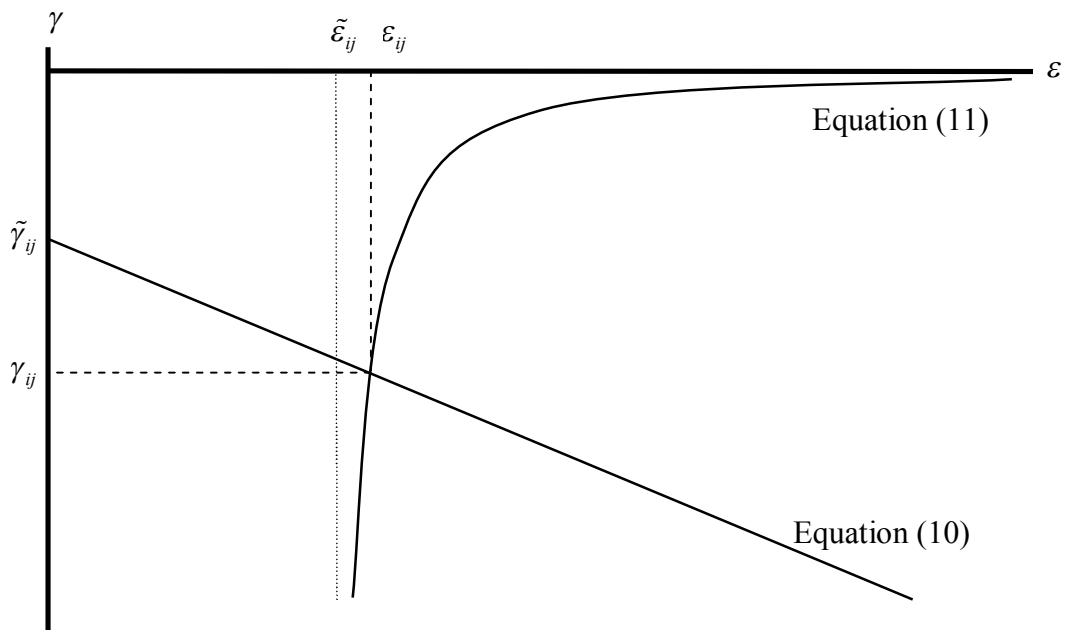**Solutions for Income and Price Elasticities (Inferior Good Case)**

Table 1. Food Expenditures in Rural China

Year	Expenditures (Yuan, current prices)		Engel Index	Shares in Total Food Expenditure (%)										
	(A) Total	(B) Food	(B)/(A)	Food	Grains	Fats & Oils	Meat	Seafood	Vegetables	Sugar	Alcohol	Fruits	Dairy Products	Other Food Products
1994	1016.81	598.47	0.589	100.00	35.99	6.68	16.75	2.92	9.19	0.80	3.02	2.40	0.25	22.00
1995	1310.36	768.19	0.586	100.00	39.01	6.35	17.19	2.95	9.07	0.97	2.01	2.37	0.23	19.84
1996	1572.08	885.49	0.563	100.00	35.29	5.33	17.83	2.94	9.65	0.92	2.01	2.65	0.27	23.10
1997	1617.15	890.28	0.551	100.00	30.84	5.49	19.11	3.00	9.65	0.87	2.05	3.66	0.33	24.99
1998	1590.33	849.64	0.534	100.00	31.14	5.61	18.55	2.86	10.39	0.89	2.08	3.17	0.33	24.98
1999	1577.42	829.02	0.526	100.00	30.76	5.43	18.27	2.83	10.91	0.88	2.10	3.18	0.34	25.29
2000	1670.13	820.52	0.491	100.00	27.70	5.52	18.51	2.96	11.93	0.65	2.11	2.97	0.38	27.27
2001	1741.09	830.72	0.477	100.00	26.02	5.04	19.43	3.00	11.51	0.76	2.09	3.22	0.43	28.50
2002	1834.31	848.35	0.462	100.00	24.79	5.19	19.30	2.98	11.28	0.81	2.15	3.00	0.40	30.09
2003	1943.30	886.00	0.456	100.00	22.65	4.72	20.63	3.04	11.90	0.54	2.11	2.79	0.55	31.08

Note: Other food products include seasonings, beans, eggs, cakes, candies, tobacco, food service, and dining out expenditures.

Source: Based on *Rural Household Survey Statistics* (various editions).

Table 2. Hedonic Model Results

		Grains		Fats & Oils		Meat		Seafood		Vegetables		Sugar		Alcohol		Fruits		Dairy Products	
		FE	RE	FE	RE	FE	RE	FE	RE	FE	RE	FE	RE	FE	RE	FE	RE	FE	RE
$\ln(PCI)$	Estimate	0.306	0.264	0.252	0.188	0.084	0.206	0.172	0.173	−0.076	0.352	−0.116	0.038	−0.007	0.009	−0.183	−0.110	0.182	0.190
	t-ratio	3.52**	3.44**	2.75**	2.55**	1.17	3.77**	2.94**	3.08**	−0.31	2.19*	−1.55	0.82	−0.15	0.21	−2.52**	−1.60	3.50**	3.61**
$\ln(HHSIZE)$	Estimate	6.490	6.334	−1.070	0.785	0.641	0.820	−8.217	−7.871	3.064	4.606	−0.871	−1.834	−3.745	−3.740	−3.630	−3.761	−1.495	−1.482
	t-ratio	3.63**	3.76**	−0.57	0.45	0.44	0.61	−6.85**	−6.73**	0.62	1.07	−0.57	−1.43	−4.00**	−3.99**	−2.43*	−2.58**	−1.40	−1.36
$\ln(HHSIZE)^2$	Estimate	−2.113	−2.019	0.078	−0.271	−0.132	−0.109	2.396	2.323	−1.524	−1.364	0.296	0.561	0.938	0.925	0.881	0.952	0.033	0.033
	t-ratio	−3.96**	−3.89**	0.14	−0.50	−0.30	−0.26	6.68**	6.60**	−1.03	−1.00	0.64	1.37	3.35**	3.28**	1.97*	2.15*	0.10	0.10
$\ln(LABOR)$	Estimate	0.349	0.286	0.867	0.581	0.191	0.200	0.082	0.049	0.966	0.976	0.133	0.256	0.527	0.552	0.928	0.985	0.652	0.666
	t-ratio	1.63	1.44	3.84**	2.87**	1.09	1.27	0.57	0.35	1.63	1.93	0.73	1.68	4.71**	4.95**	5.19**	5.72**	5.12**	5.13**
$\ln(LAND)$	Estimate	0.142	0.102	0.006	−0.004	0.075	0.049	−0.070	−0.064	0.179	−0.066	0.118	0.110	0.017	0.016	−0.039	−0.05	0.056	0.046
	t-ratio	2.94**	2.50**	0.12	−0.09	1.89	1.73	−2.14*	−2.07*	1.33	−0.80	2.84**	4.62**	0.66	0.63	−0.95	−1.47	1.95*	1.57
$\ln(HOUSE)$	Estimate	−0.045	−0.044	0.037	0.036	0.065	0.116	−0.164	−0.164	−0.043	0.154	0.049	0.045	0.035	0.038	0.062	0.076	−0.114	−0.107
	t-ratio	−0.61	−0.62	0.47	0.48	1.08	1.96*	−3.33**	−3.38**	−0.21	0.81	0.77	0.80	0.91	0.99	1.01	1.24	−2.59**	−2.39*
$\ln(EDEXP)$	Estimate	−0.045	−0.046	0.014	0.012	−0.016	−0.007	0.003	0.003	−0.067	−0.058	−0.031	−0.023	0.006	0.007	0.002	0.004	−0.030	−0.029
	t-ratio	−4.68**	−4.86**	1.41	1.17	−1.98*	−0.86	0.49	0.50	−2.52**	−2.25*	−3.81**	−2.93**	1.11	1.31	0.25	0.53	−5.17**	−5.00**
$EDLEVEL$	Estimate	−0.656	−0.377	−0.715	−0.310	−0.310	0.064	−0.235	−0.226	0.517	0.470	0.476	0.339	−0.084	−0.064	−0.200	−0.299	0.172	0.191
	t-ratio	−2.26*	−1.45	−2.33*	−1.25	−1.30	0.35	−1.20	−1.20	0.64	0.88	1.90	2.18*	−0.55	−0.42	−0.82	−1.29	0.99	1.08
t	Estimate	−0.026	−0.027	−0.059	−0.053	−0.013	−0.026	−0.048	−0.047	0.017	0.007	−0.015	−0.024	−0.019	−0.021	−0.013	−0.015	−0.019	−0.020
	t-ratio	−3.86**	−5.14**	−8.43**	−11.16**	−2.37**	−7.42**	−10.84**	−11.22**	0.90	0.76	−2.64**	−8.38**	−5.57**	−6.30**	−2.32*	−3.03**	−4.70**	−4.92**
Intercept	Estimate	−7.055	−6.776	0.947	−0.498	0.581	−0.957	7.993	7.652	−2.186	−8.388	2.306	1.996	3.619	3.480	3.844	3.309	2.196	2.082
	t-ratio	−4.25**	−4.30**	0.54	−0.31	0.43	−0.77	7.17**	7.01**	−0.47	−2.16*	1.62	1.73	4.16**	3.98**	2.77**	2.43*	2.22*	2.05*
R^2		0.726	0.724	0.809	0.804	0.536	0.512	0.842	0.842	0.157	0.106	0.548	0.538	0.504	0.503	0.476	0.472	0.393	0.392
Hausman test: fixed effects vs. random effects		33.65		14.22		24.46		2.73		13.24		25.59		19.27		10.98		−45.83	
p-value		0.0001		0.1146		0.0036		0.9742		0.1521		0.0024		0.0230		0.2774		NA	

Notes: FE = fixed effects model, RE = random effects model.   NA signifies not applicable because the results for dairy products fail to satisfy the asymptotic assumptions of the Hausman test (the test statistic is negative).   All R^2 values are within-group results.

** Significant at 1% level.   * Significant at 5% level.

Table 3. Elasticity Corrections for Recent Demand Studies

	Grain		Vegetables		Seafood		Fats & Oils
	Huang and Rozelle (1998)	Shen (2001)	Huang and Rozelle (1998)	Shen (2001)	Ma et al. (2004)	Shen (2001)	Shen (2001)
Study Estimates							
Income Elasticity	0.510	0.105	1.400	0.598	1.280	1.638	0.851
Price Elasticity	-0.570	-0.147	-0.820	-0.154	-0.521	-0.925	-0.130
After Correction							
Income Elasticity	0.389	0.077	1.169	0.549	1.197	1.493	0.827
Price Elasticity	-0.395	-0.093	-0.658	-0.140	-0.485	-0.835	-0.126
Overestimation (%)							
Income Elasticity	31	37	20	9	7	10	3
Absolute Value of Price Elasticity	44	59	25	10	8	11	3

Footnotes

---

[1] A parallel study for urban China would be highly valuable but data availability and quality are limitations. The China National Statistics Bureau (CNSB) does not publish data on unit values for urban China. Also, dining out expenditures, which are significantly greater in urban areas than in rural areas, are not captured well by CNSB data.

[2] Q_{ij} is an accounting measure of the quantity consumed of a food group as commonly found in household surveys and should not be confused with the group quantity in a multistage demand system (e.g. Deaton and Muellbauer 1980, Moschini 2001).

[3] An alternative type of separability is indirect weak separability in which the indirect utility function depends on indices for each group. Each group index depends on prices of goods within that group and total expenditure (Moschini 2001).

[4] This holds for small changes in prices within a food group. Large changes in prices could cause a switch in which item is the least expensive within a group, leading to a movement from one corner solution to another.

[5] As Deaton (1988) indicates, one would expect $\varepsilon_{ij}\eta_{ij} < |\gamma_{ij}|$, so that $0 < 1 - \varepsilon_{ij}\eta_{ij}/\gamma_{ij} < 1$ when $\gamma_{ij} > 0$ and $1 < 1 - \varepsilon_{ij}\eta_{ij}/\gamma_{ij} < 2$ when $\gamma_{ij} < 0$.

[6] A *mu* is a traditional Chinese measure of land area, with 15 *mu* equal to one hectare.

[7] To illustrate the calculations involved in arriving at the corrected price and income elasticities of demand, consider the case of grains and drop the subscripts i and j for ease of exposition. Huang and Rozelle (1998) find that $\tilde{\gamma} = 0.510$ and $\tilde{\varepsilon} = 0.570$ for grains, and our results indicate that $\eta = 0.306$. Then the b and c terms in equation (12a) are $b = 0.510 - 2(0.570)(0.306) = 0.161$ and $c = (0.570)(0.306)(0.510) = 0.089$. Using equation

(12a), $\gamma = \left(0.161 + \sqrt{(0.161)^2 + 4(0.089)}\right)\Big/2 = 0.389$.   Using equation (13),

$\varepsilon = \left(0.510 - 0.389\right)\big/0.306 = 0.395$.